WHAT is missing

Overall stuff

Write introductory text

Attempt at introduction:
We are four students, who created this dashboard during a course in data visualization in an attempt to illustrate/find different relations between Major Security and Safety Events in transportation in the United States of America between 2014 and 2022, which is our dataset. The dataset can be found at:
Link
<link_to_dataset>(https://catalog.data.gov/dataset/major-safety-events).

The dashboard includes fully interactive plots using Plotly. This means that all plots can be adjusted to highlight certain areas or specific information. The plots allow zoom in and out. These features are incredibly useful on the heat maps if specific areas are interesting for the user.

At the top of the page a navigation bar can be found. The navigation bar can be used to switch between the different pages. Each page contains an overall question, which we tried to answer through the visualizations. The questions are the following:

1. Is there an increase or decrease in certain types of accidents in the last 9 years?
2. Is there a relation between time periods in the day and certain types of accidents?
3. Do certain types of accidents occur more often in certain environments?

The report can be downloaded here.

At Least 8 graphs in total

- must be a minimum of 3 different types
- at least 1 animated graph

NOTE: Add indications. Why are collisions lower in 2020

Questions

1. Is there a relation between time periods in the day and certain types of accidents?
Current:

- We have a 3d graph

To Do:

- Should be made into multiple 2d graphs showing the same content.
- Text about what the visualization is about and trying to show

2. Is there an increase or decrease in certain types of accidents in the last 9 years?

Current:

- We have a bar chart and linechart

To Do:

- 2023 should be removed since we don't care
- bar chart should perhaps be showed in lowest to highest order
- Write some text about the goal of what we are trying to show
- Add indications explaining why the dataset dips in 2019

3. Do certain types of accidents occur more often in certain environments?

Current:

- We have population data from 2014-2023, so the states/counties can be correctly showed

To Do:

- Create a heat map with states or counties with event types as a selection bar. The counties/states should be normalized so it is accidents / xxxx capita. (year could be an additional identifier)
- Perhaps do a side by side with an ordered horizontal bar chart showing the accidents of each county / state in different years. The bar chart/map could be animated to go from 2014-jan to 2022-dec.

After stuff

Create a link to download report from our dash app
Perhaps make the text interactive, so the text could be: As seen in collisions the amount of collisions dropped in 2019. When the “collisions” text is clicked on the graph is updated to visualize collisions.

Data Visualization - Group 11

Tobias Vittrup Bak - tobib21@student.sdu.dk
Thor Malmby Jørgin - tjoer21@student.sdu.dk
Kevin Torp - ketor21@student.sdu.dk
Philip Schwartz - pschw21@student.sdu.dk

Abstract

Briefly describe the data and what were the main achievements of your visualizations

Table of Contents

1. Background and Motivation

Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

The dataset seemed relevant and interesting since it immensely impacted the users in the dataset.

2. Project Objectives

Provide a list of questions (and sub questions if relevant) which you plan to answer with your visualizations. What would you like to learn and accomplish?

4. Is there a relation between time periods in the day and certain types of accidents?
5. Is there an increase or decrease in certain types of accidents in the last 9 years?
6. Do certain types of accidents occur more often in certain environments?
7.

We would like to learn and explore how a dataset of this scale can be effectively visualized for easier information gathering by users.

3. Data

- From where did you get your data? If appropriate, provide a link to your data sources.
- Description: describe all the relevant variables, number of records and any special feature of your data (if there is any)
- Data Processing. Did you need to do substantial data cleanup? If yes, what techniques did you use?

The data was collected by the U. S. Government. Data origin: https://catalog.data.gov/dataset/major-safety-events

The dataset contains entries for major safety events that have happened in the United States, in the period 2014-2023. There are 82916 records. The important variables are:

- **Time, time 2, time 3:** These three variables relatively record the day and time, which is the exact time the accident/major safety events occurred.

Data processing:
We removed some columns, which were completely empty, such as hazardous. We removed only columns that we would not be using for the visualizations or the parsing either. The dataset was somewhat messy, since it consists of aggregated data from several counties. Therefore there were more than 100 columns, each with small details, that only some counties used, and only for some types of safety events. DATA TECHNIQUE??

4. Visualization/Dashboard

- **Design**. How will you display your data? Provide some general ideas that you have for the visualization design. Describe your designs and justify your choices of visual encodings.
- **Must-Have Features**. List the features without which you would consider your project to be a failure.
    - 1. You must have at least three types of graphs (i.e barchart, timeseries plot or boxplots) and
    - 2. at least one animated graph (using for example gganimate).
    - 3. In total at least 8 graphs. Provide clear and well-referenced images showing the key design and interaction elements.
    - 4. A link to the dashboard/Visualization must also be included in the report.
    - 5. An option to download the report as a manual from the dashboard
- **Optional Features.** List the features which you consider to be nice to have, but not critical.

5. Story/Results

here you should provide answers to the questions in 2.Project Objectives. Tell the story of the data that you saw in the visualization. What were your expectations and how close they were to what data revealed. What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

6. Conclusion/Discussion

Finally, you conclude the report by a summary of what you achieved, how you achieved, what were the challenges for you and how the course can be improved.

7. Participation

Student	PoC
Tobias Vittrup Bak, tobib21@student.sdu.dk	1000
Thor Malmby Jørgin, tjoer21@student.sdu.dk	1000
Kevin Torp, ktorp@student.sdu.dk	1000
Philip Schwartz, pschw21@student.sdu.dk	1000

References:

USA STATE POPULATION DATA:
https://usafacts.org/data/topics/people-society/population-and-demographics/population-data/population/

USA COUNTIES POPULATION DATA:
https://www2.census.gov/programs-surveys/popest/datasets/